

# Modélisation probabiliste en biologie

Florent Malrieu & Romain Yvinec

16 Janvier 2020

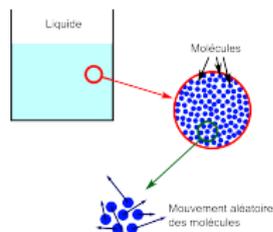
# Déterministes vs Stochastiques

✓ La plupart des systèmes sont soumis à des **perturbations**, externes ou internes, qui peuvent être très compliquées à représenter finement et précisément, et pour lesquelles la formalisation en tant que perturbation **aléatoire** est très pratique.



# Déterministes vs Stochastiques

- ✓ La plupart des systèmes sont soumis à des **perturbations**, externes ou internes, qui peuvent être très compliquées à représenter finement et précisément, et pour lesquelles la formalisation en tant que perturbation **aléatoire** est très pratique.
- ✓ Dans certains modèles, le caractère aléatoire est justifié pour représenter des phénomènes physiques qui reposent sur des **interactions trop complexes** pour être modélisées. (Le mouvement d'une particule en solution, bien que régit par les équations de la physique (déterministes), est très représenté par un mouvement *Brownien* (aléatoire).



# Déterministes vs Stochastiques

- ✓ La plupart des systèmes sont soumis à des **perturbations**, externes ou internes, qui peuvent être très compliquées à représenter finement et précisément, et pour lesquelles la formalisation en tant que perturbation **aléatoire** est très pratique.
- ✓ L'occurrence d'événement **ponctuels**, qui peuvent dépendre de nombreux autres événements, est aussi souvent modélisée de manière aléatoire (radioactivité, arrivée d'un client dans une file d'attente, passage de l'autobus, phénomène météorologique...).



# Déterministes vs Stochastiques

- ✓ La plupart des systèmes sont soumis à des **perturbations**, externes ou internes, qui peuvent être très compliquées à représenter finement et précisément, et pour lesquelles la formalisation en tant que perturbation **aléatoire** est très pratique.
- ✓ Certains systèmes sont enfin **intrinsèquement** aléatoires (mutation génétique, lancé de dés..)



# Déterministes vs Stochastiques en biologie

- ✓ En biologie de l'évolution : L'aléatoire est au coeur de la théorie de l'évolution pour expliquer l'émergence de nouveaux phénotypes à partir de mutations génétiques aléatoires.
- ✓ En dynamique des populations : la modélisation aléatoire est pertinente notamment lorsqu'on a de petits effectifs, ou que la variabilité inter-individuelle est importante.

# Déterministes vs Stochastiques en biologie

- ✓ En biologie moléculaire : il y a de nombreux exemples de phénomènes (apparemment) aléatoires au niveau microscopique, due à des mouvements aléatoires et/ou de petits effectifs :
- Fluctuations thermiques de réactions biochimiques
  - Division cellulaire asymétrique
  - Sélection aléatoire des gamètes et du matériel génétique
  - Système immunitaire : création de nouveaux anti-corps
  - Système nerveux : activité électrique modulée par des réactions microscopiques
  - Expression des gènes : le lien génotype->phénotype n'est pas univoque.

✓ Contrairement à la modélisation déterministe, deux simulations du même modèle, avec les mêmes paramètres, ne donneront pas des trajectoires identiques.

⇒ On ne s'intéresse pas en général à une seule trajectoire mais à des **statistiques** liées au modèle : moyenne, variance, et plus globalement la *loi de probabilité* (ou distribution de probabilité).

✓ Contrairement à la modélisation déterministe, deux simulations du même modèle, avec les mêmes paramètres, ne donneront pas des trajectoires identiques.

⇒ On ne s'intéresse pas en général à une seule trajectoire mais à des **statistiques** liées au modèle : moyenne, variance, et plus globalement la *loi de probabilité* (ou distribution de probabilité).



**Aléatoire**  $\neq$  **imprévisible**. ⇒ Très souvent, toutes les sorties d'un modèle ne sont pas équiprobables : il y a des évènements plus probables que d'autres, et des phénomènes déterministes sous-jacents à des observations aléatoires.

✓ Comme pour la modélisation déterministe, il y a plusieurs classes de modèles :

- temps discret / temps continu
- espace d'états discrets / espaces d'états continus.

⇒ On se limitera aujourd'hui aux espaces d'états discrets (ex : nombre d'individus dans une population). Les espaces d'états continus sont en particulier bien adapté pour modéliser des mouvements spatiaux aléatoires (ex : mouvement brownien d'une particule).

✓ On distingue deux types d'approches numériques : des simulations répétées de trajectoires ("aléatoire") ou la simulation de l'évolution des lois de probabilité (qui sont régis par des équations déterministes!).

⇒ On abordera aujourd'hui seulement les simulations de trajectoires, en général plus simple à comprendre et à simuler.

✓ La notion d'**équilibre** est très différente d'un équilibre déterministe. Un état est **stationnaire** dans un modèle probabiliste si sa loi de probabilité est constante au cours du temps (mais les trajectoires peuvent continuer à évoluer).

✓ La notion d'**équilibre** est très différente d'un équilibre déterministe. Un état est **stationnaire** dans un modèle probabiliste si sa loi de probabilité est constante au cours du temps (mais les trajectoires peuvent continuer à évoluer).

⇒ Heuristiquement, on observe typiquement plusieurs types de comportements pour les trajectoires en temps long :

- La trajectoire n'évolue plus dans le cas d'un point absorbant (*ex : taille de la population égale à 0 en dynamique de population*)
- La trajectoire "fluctue faiblement" autour d'un point (*typiquement l'équilibre déterministe du modèle sous-jacent*) ou d'une trajectoire périodique (*cas d'oscillations comme dans le modèle de Lotka-Volterra*)
- La trajectoire effectue de petites "fluctuations" entre des "grands sauts" (*typiquement présence de plusieurs états stationnaires stables pour le modèle déterministe sous-jacent*)

# Chaîne de Markov

Les modèles probabilistes de type chaîne de Markov sont définies par les ingrédients essentiels suivants :

- Une liste d'évènement possibles
- Un taux de fréquence moyen d'apparition pour chacun de ces évènements (en temps continu) ou une probabilité d'apparition pour chacun de ces évènements (en temps discret)

Ce principe permet de construire des algorithmes très simples pour simuler des trajectoires et étudier l'évolution de différentes statistiques au cours du temps

# Marche aléatoire

La marche aléatoire en dimension 1 est le modèle "de base" pour avoir de l'intuition sur les types de comportements possibles :

- "divergence" vers l'infini
- occupation uniforme de l'espace (condition réfléchi)
- absorption en un point (condition absorbante)
- "petites fluctuations" autour d'un point (potentiel confinant)
- "petites fluctuations" entre coupée par des "grands sauts" autour de différents points (potentiel multistable)

Ce modèle a de nombreux liens avec les modèles déterministes ODE et EDP.

# Modèle de dynamique de population

On a vu l'équivalent stochastique des modèles de dynamique de populations :

- croissance exponentielle de Malthus (Galton-Watson)
- Modèle logistique
- Modèle à effet Allee
- Modèle de compétition de Lotka-Volterra

La principale différence entre la modélisation déterministe et stochastique en dynamique des populations réside dans le fait que la possibilité d'extinction est toujours possible dans les modèles stochastiques (0 est absorbant).

# Autre modèles en biochimie

Sur le principe des chaînes de Markov, on peut construire des modèles stochastiques associés à n'importe quel réseau de réactions biochimiques.

- A chaque réaction correspond un événement (une transition) possible.

On a vu comme exemple le modèle d'expression des gènes (ANR, Protéine).